J Mol Model (2009) 15:357–367
DOI 10.1007/s00894-008-0402-0

ORIGINAL PAPER

Computational studies of the binding mode and 3D-QSAR analyses of symmetric formimidoester disulfides: a new class of non-nucleoside HIV-1 reverse transcriptase inhibitor

Elena Cichero · Sara Cesarini · Andrea Spallarossa · Luisa Mosti · Paola Fossa

Received: 31 July 2008 / Accepted: 16 September 2008 / Published online: 9 December 2008

Abstract Symmetric formimidoester disulfides (DSs) have recently been identified as a new class of potent non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors. Given that three geometric isomers for DSs are possible, a computational strategy based on molecular docking studies, followed by comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) was used in order to identify the most probable DS isomer interacting with RT, to elucidate the atomic details of the RT/DS interaction, and to identify key features impacting DS antiretroviral activity. The CoMFA model was found to be the more predictive, with values of $r^2_{ncv} = 0.95$, $r^2_{cv} = 0.482$, SEE=0.264, F=80, and $r^2_{pred} = 0.73$.

Keywords CoMFA · CoMSIA · 3D-QSAR · Disulfides · Docking · HIV-1 · Non-nucleoside reverse transcriptase inhibitors

Introduction

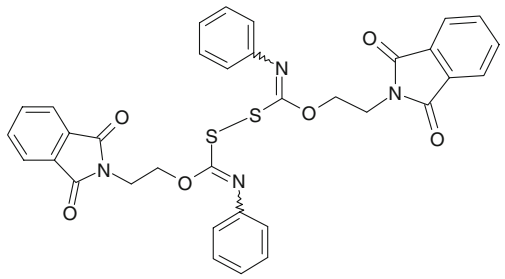
Reverse transcriptase (RT) is a key enzyme in the HIV replication cycle and is one of the main targets in the development of drugs for treating HIV-infection and AIDS [1–5]. RT catalyses the conversion of viral RNA into double stranded DNA, which is then integrated in the host genome. Non-nucleoside RT inhibitors (NNRTIs) bind to an allosteric hydrophobic pocket (NNRTI binding site, NNIBS), located at about 10 Å from the polymerase active site. Upon inhibitor interaction, the NNIBS is created, locking the enzyme into an inactive form by affecting the geometry of the polymerase active site aspartyl residues [6]. In the past 15 years, more than 50 structurally diverse NNRTIs have been described [6–12]. The fact that cross-resistance extends to the whole NNRTI class calls for the development of new agents capable of inhibiting clinically relevant NNRTI-resistant mutants [13, 14].

In previous studies, we reported the discovery of the potent NNRTI class of *O*–(2-phthalimidoethyl)–*N*-arylthiocarbamates (C-TCs) [15, 16] and structurally related compounds, such as ring-opened analogues (O-TCs) [15, 16], *N*-acylated derivatives (ATCs) [17, 18] and non-phthalimidic congeners (TCs) [19, 20]. More recently, the molecular duplication of the isothiocarbamic form of C-TCs led to symmetric formimidoester disulfides (DS; Fig. 1), which were subsequently identified as a novel class of potent NNRTI [21].

In order to elucidate the molecular basis of RT/DS interactions (in the absence of crystallographic data for the RT/DS complex), we performed docking studies on a series of 29 DS compounds using the X-ray structure of RT in complex with C-TC *O*-[2-(phthalimido)ethyl]-*N*-(4-chlorophenyl)thiocarbamate (**I**, Fig. 2) [22]. Our aim was also to identify features significantly impacting DS antiretroviral activity, to elaborate a quantitative structure-activity relationship (QSAR) model, and to obtain useful suggestions for the design of new DSs with improved potency, also against clinically relevant resistant mutants.

Thus, DS **1–29** were considered in the three possible geometric isomers (Fig. 3), generating three databases including *E*,*E*; *E*,*Z* and *Z*,*Z* isomers, respectively, and treated as three different sets for docking simulation. All the best ranked docking poses, selected from any database, were then compared according to the docking final scoring

E. Cichero (✉) · S. Cesarini · A. Spallarossa · L. Mosti · P. Fossa
Dipartimento di Scienze Farmaceutiche,
Università degli Studi di Genova,
Viale Benedetto XV n.3,
16132 Genoa, Italy
e-mail: cichero@unige.it



1

Fig. 1 Chemical structure of DS **1**

function (S), in order to identify the most probable DS isomer interacting with RT (the one showing the lowest mean S value). Subsequently, 3D-QSAR studies involving comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA) were performed on DS **1–29** in the configuration that had emerged as the most probable from the first part of the analysis.

Methods

Data set

A dataset of 29 disulfides (DSs) (Table 1), screened according to the same pharmacological protocol, were selected from our compound collection [21]. The *E*,*E*; *E*,*Z* and *Z*,*Z*, isomers of **1**–**29** were built, parameterised (Gasteiger-Huckel method) and energy minimised within MOE using MMFF94 forcefield [23].

Docking protocol

Starting from three databases, comprising the *E*,*E*; *E*,*Z* and *Z*,*Z* isomers of **1**–**29**, a docking procedure was performed. In absence of crystallographic data for an RT/DS complex, the three-dimensional (3D) structure co-ordinates of RT in complex with C-TC (**I**, Fig. 2) (PDB entry 2VG5) was used as a starting point for a preliminary manual docking simulation [22]. In more detail, one DS (compound **13**) with a high pEC_{50} value of was selected as a reference compound. Starting from its three isomers, the corresponding *E*,*E*; *E*,*Z* and *Z*,*Z* RT/**13** complex models were derived by superimposition of each isomer of **13** on the RT/**I** crystallographic structure (manual docking). Thereafter, an automated docking procedure of the initial three databases, including the respective *E*,*E*; *E*,*Z* and *Z*,*Z* isomers of **1-29** was performed using the corresponding previously derived *E*,*E*; *E*,*Z* and *Z*,*Z* RT/**13** complexes as ligand/receptor model.

Manual docking

The three isomers of compound **13**, one of the most active of the series, were manually superimposed on the crystallographic structure of **I** in complex with RT, according to the common *O*-[2-(phthalimido)ethyl]-*N*-phenylthiocarbamate moiety, generating three DS/RT complexes (**I** was erased after DS superimposition).

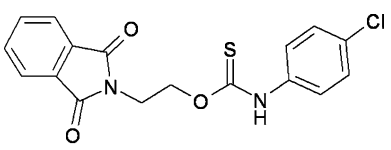
Each complex was minimised within LigX, a module of the MOE software. Briefly, receptor atoms far from the ligand were held fixed (constrained not to move at all), while residues within a certain distance (8 Å) could move so that ligand atoms were not fixed. A forcefield (MMFF94) energy minimisation was performed, which terminated when the root mean square (RMS) gradient of potential energy fell below a certain threshold, set to 0.05 kcal/molÅ in this case. A value of the MM/GBVI scoring function, related to ligand and receptor Van der Waals, Coulomb and generalised Born implicit solvent interactions energy, was associated to each complex. Each of the three DS/RT final complexes (RT in complex with *E*,*E*; *E*,*Z* or *Z*, *Z* **13**) was employed for the automated molecular docking simulation of the corresponding DS database (*E*,*E*; *E*,*Z* or *Z*,*Z* **1-29**).

Automated docking

Each geometric isomer was docked into the NNIBS using the flexible docking module implemented in MOE. For all compounds, the best-docked geometries, evaluated in terms of "London dG", were refined by energy minimisation (MMFF94) and rescored according to “Affinity dG” (kcal/mol of total estimated binding energy). Following this procedure, on the basis of the final docking scoring function (S), we identified the most probable DS geometric isomer (*E*,*Z*) interacting with RT (lowest mean S value).

3D-QSAR analysis

Although a structure-based approach for DS/RT interaction analyses on the basis of the RT/**I** complex is possible, the ligand-based approach of CoMFA and CoMSIA analyses



I

Fig. 2 Chemical structure of C-TC {*O*-[2-(phthalimido)ethyl]-*N*-(4-chlorophenyl)thiocarbamate} **I**

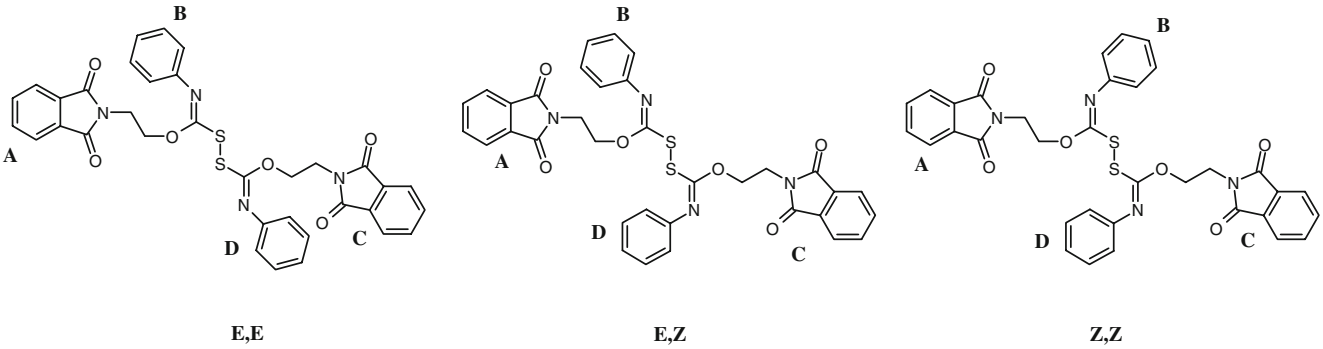


Fig. 3 The three possible geometric isomers of **1** used in the computational analysis

[24, 25] could provide a complementary tool for drug design.

The prior analysis revealed the *E,Z* isomers of **1-29** as the most probable isomers to interact with RT; these were then manually aligned on the basis of the common disulfide moiety, in order to directly develop CoMFA and CoMSIA analyses using Syby7.0 software [26].

Training set and test set

All the compounds were grouped into a training set, for model generation, and a test set, for model validation, containing 23 and 6 compounds, respectively. The molecules of the test set represent 20% (considered an appropriate percentage to validate a molecular model) of the training set. Both the training and the test set were divided manually according to a representative range of biological activities and structural variations. For QSAR analysis, EC_{50} values were transformed into pEC_{50} values and then used as response variables. The enzyme-inhibitory activity of these compounds covered 4 log orders of magnitude.

CoMFA and CoMSIA analysis

CoMFA is a widely used 3D-QSAR technique that relates the biological activity of a series of molecules with their steric and electrostatic fields. The latter are calculated by placing the aligned molecules, one by one, into a 3D cubic lattice with a 2 Å grid spacing. The van der Waals potential and Coulombic terms, which represent steric and electrostatic fields, respectively, were calculated using the standard Tripos force field method. The column-filtering threshold value was set to 2.0 kcal/mol to improve the signal-to-noise ratio. A methyl probe with a +1 charge was used to calculate the CoMFA steric and electrostatic fields. A 30 Kcal/mol energy cut-off was applied to avoid infinity of energy values inside the molecule.

The CoMSIA method calculates five descriptors, namely steric, electrostatic, and hydrophobic parameters and the H-bond donor and H-bond acceptor properties. The similarity indices descriptors were calculated using the same lattice box employed for CoMFA calculations, using an sp^3 carbon as probe atom with a +1 charge, +1 hydrophobicity and +1 H-bond donor and +1 H-bond acceptor properties.

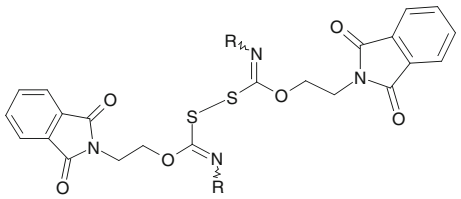
Partial least square analysis and models validation

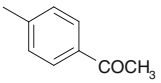
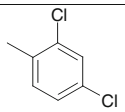
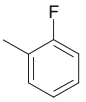
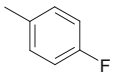
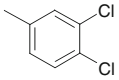
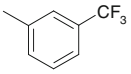
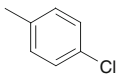
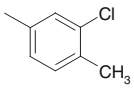
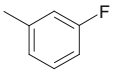
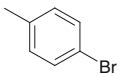
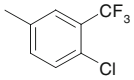
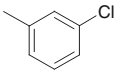
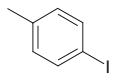
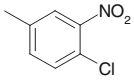
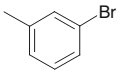
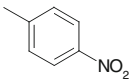
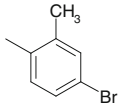
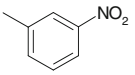
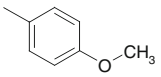
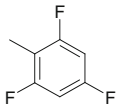
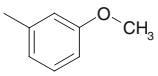
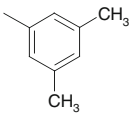
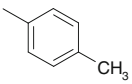
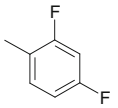
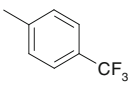
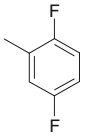
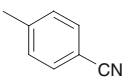
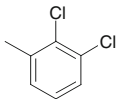
The partial least-squares (PLS) approach, an extension of the multiple regression analysis, was used to derive the 3D-QSAR models. CoMFA and CoMSIA descriptors were used as independent variables and pEC_{50} values were used as dependent variables. Prior to PLS analysis, CoMFA and CoMSIA columns with a variance of less than 2.0 kcal mol^{-1} were filtered by using column filtering to improve the signal-to-noise ratio.

The leave one out (LOO) cross-validation method was used to check the predictivity of the derived model and to identify the optimal number of components (ONC) leading to the highest cross-validated r^2 $\left(r^2_{cv}\right)$. In the LOO methodology, one molecule is omitted from the dataset and a model involving the rest of the compounds is derived. Employing this model, the activity of the omitted molecule is then predicted.

The ONC obtained from cross-validation methodology was used in the subsequent regression model. Final CoMFA and CoMSIA models were generated using non-cross-validated PLS analysis. To further assess the statistical confidence and robustness of the derived models, a 100-cycle bootstrap analysis was performed. This is a procedure in which *n* random selections out of the original set of *n* objects are performed several times (100 times were required to obtain good statistical information). In each run, some objects may not be included in the PLS analysis, whereas some others might be included more then once. The mean correlation coefficient is represented as bootstrap

Table 1 Molecular structure of disulfides (DS) **1–29**



Comp.	R	Comp.	R	Comp.	R
1		**12**		**23**	
2		**13**		**24**	
3		**14**		**25**	
4		**15**		**26**	
5		**16**		**27**	
6		**17**		**28**	
7		**18**		**29**	
8		**19**			
9		**20**			
10		**21**			
11		**22**			

$r^2\left(r^2_{boot}\right)$. To validate the CoMFA- and CoMSIA-derived models, the predictive ability for the test set of compounds (expressed as r^2_{pred}) was determined by using the following equation:

$$r^2_{pred} = (SD - PRESS)/SD$$

SD is the sum of the squared deviations between the biological activities of the test set molecules and the mean activity of the training set compounds. PRESS is the sum of the squared deviation between the observed and the predicted activities of the test set compounds.

All calculations were carried out using a PC running the Windows XP operating system and a PC running the Linux Red Hat operating system.

Results and discussion

Docking protocol

Manual docking

According to our calculations, the *E*,*E* and *E*,*Z* isomers of **13**, superimposed on the crystallographic structure of **I** in complex with RT (Fig. 4), display an H-bond interaction between the K101 backbone NH and the nitrogen atom of the isothiocarbamic function, located in the inner part of the NNIBS cavity, while the *Z*,*Z* isomer displays no H-bonds.

Moreover, the *E*,*Z* isomer shows an H-bond between the K103 ε-amino group and one of the carbonyl oxygens of the phthalimide moiety located outside the NNIBS.

On the basis of the LigX refinement calculations, the *E*,*Z* isomer is predicted to be 3.26 and 1.05 kcal/mol (MM/GBVI scoring function), respectively, lower in energy than the *Z*,*Z*- and *E*,*E*- isomers.

Automated docking

E,E geometric isomer docking poses As shown in Fig. 5a, all the *E*,*E* isomers display an H-bond interaction between the V179 backbone NH and the nitrogen atom of the isothiocarbamic function, oriented towards the external side of the NNIBS cavity. The phenyl ring located inside the NNIBS pocket is engaged in hydrophobic contacts with L100, L234, P236 and Y318. The phthalimide moiety positioned inside NNIBS shows hydrophobic contacts with P95 and π–π stacking with Y181, Y188 and W229, while the other phthalimide moiety is located outside the enzyme pocket, and establishes a cation–π interaction with the K101 ε-amino group.

E,Z geometric isomer docking poses As depicted in Fig. 5b, the phthalimide moiety C is positioned outside the NNIBS pocket and establishes an H-bond (by one of its carbonyl functions) with the K103 ε-amino group, and hydrophobic contacts with P321. In addition, the phthalimide moiety A

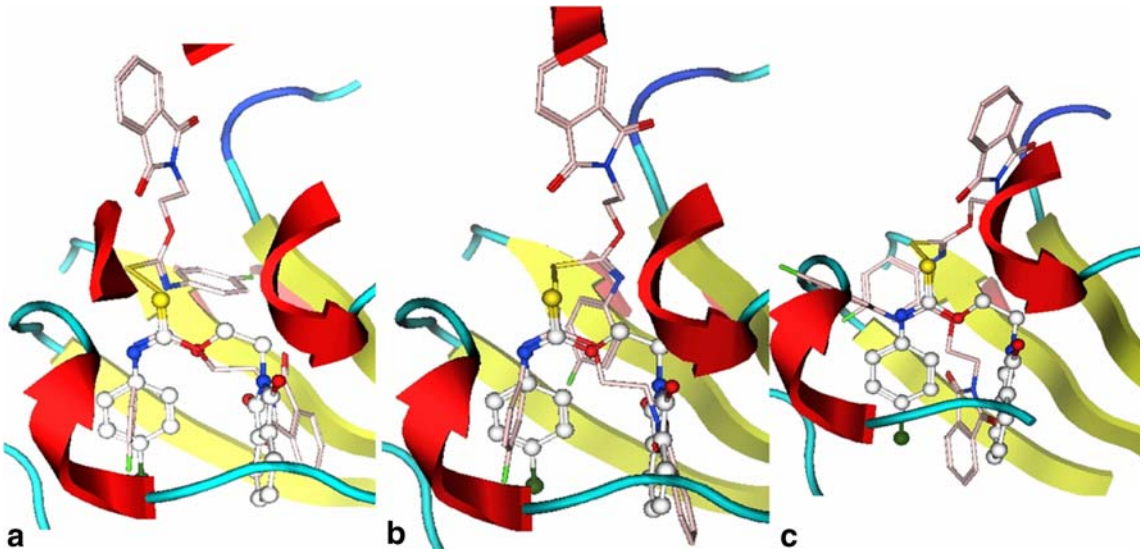


Fig. 4 Manual docking by superimposition of **13** (**a** *E*,*E* isomer; **b** *E*,*Z* isomer, **c** *Z*,*Z* isomer) on the X-ray structure of C-TC in complex with RT. C-TC is depicted in *ball and stick* representation, coloured by atom type. Compound **13** is shown in *stick mode*. Colour code: *pink* C, *blue* N, *red* O, *yellow* S

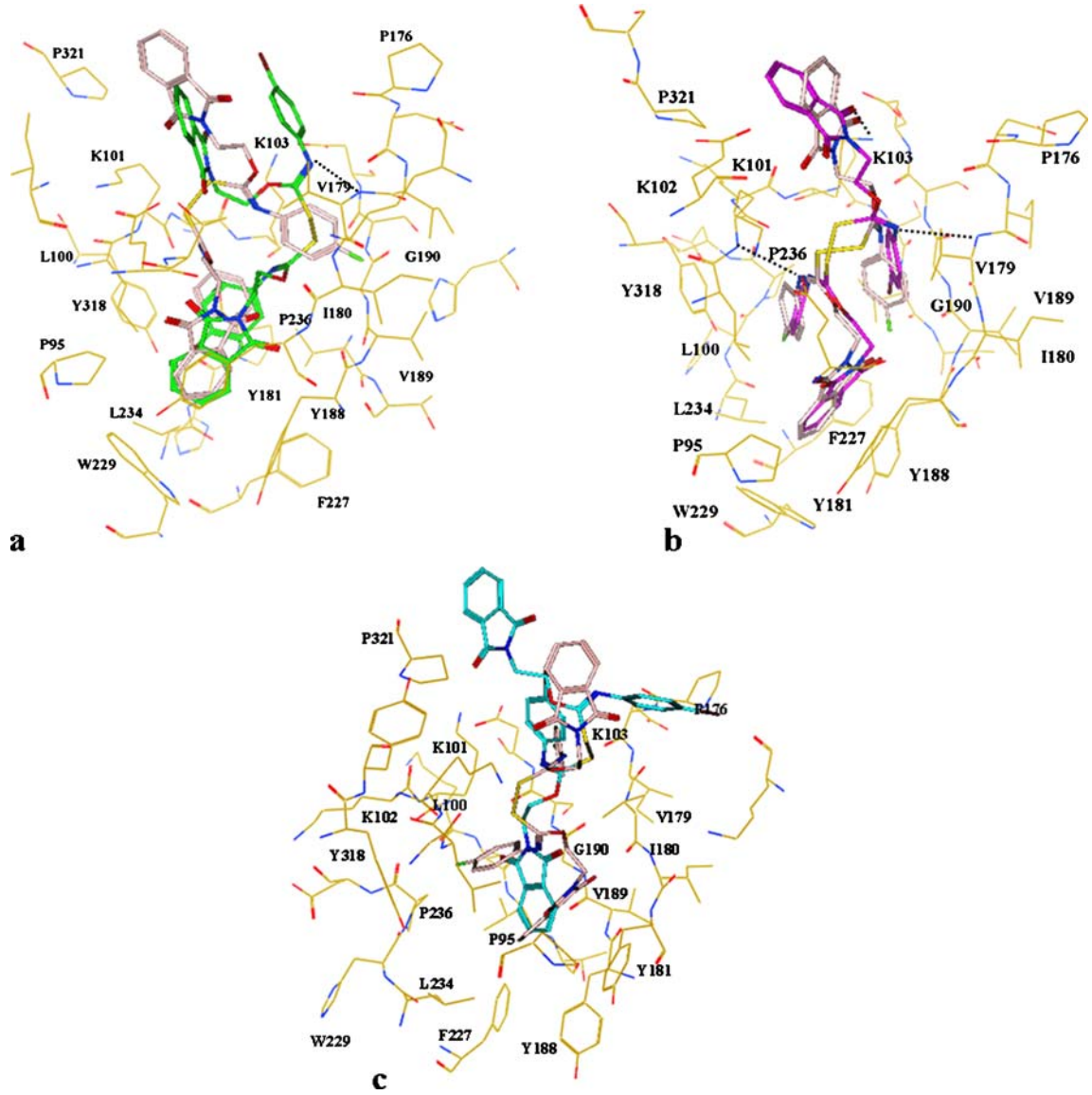


Fig. 5 Automated docking analysis of DS/RT complexes, obtained by a manual docking procedure followed by energy minimisation of the resulting complexes. Compound **13** is shown in *stick mode* (colour code: *pink* C, *blue* N, *red* O, *yellow* S). Compound **15** (**a** *E*,*E* isomer; **b** *E*,*Z* isomer; **c** *Z*,*Z* isomer) docking poses are depicted in *stick mode*. Only residues within 5 Å from ligands are reported, the most important are labelled. Hydrogen bonds are coloured *black*

is located inside the NNIBS cavity, and shows hydrophobic contacts with P95 and π–π stacking with the aromatic residues of Y181, Y188 and W229. The *N*-phenyl ring B is oriented towards a lipophilic region formed by L100, K101, L234, P236 and Y318, and establishes an H-bond with the K101 backbone NH via the nitrogen atom. The *N*-phenyl D occupies a hydrophobic pocket consisting of V179, V189, G190, displaying an H-bond interaction with the V179 backbone NH.

Z,Z geometric isomer docking poses The only ring found located inside the NNIBS is one of the phthalimides (Fig. 5c), which displays hydrophobic contacts with P95, L100, and G190, while the other rings are oriented towards the external side of the pocket. No H-bonds are detected.

Taking into account these data and the final docking scoring function for all the isomers of DS **1**–**29**, as reported in Table 2, we identify the *E*,*Z* geometric isomer as the most probable isomer interacting with RT. To verify this hypothesis, the selected docking pose of DS **15** *E*,*Z* isomer was compared with the RT surface electrostatic distribution (Connolly surface), as illustated in Fig. 6. A good correlation between the hydrophilic and hydrophobic

Table 2 Final docking scoring function for the three isomers of DS **1**–**29**

Compound	Docking Final Score (S)		
	E,*E* isomer	*E*,*Z* isomer	*Z*,*Z* isomer
1	−11,104	−11,580	−6,396
2	−11,912	−11,807	−8,202
3	−11,059	−13,111	−8,994
4	−11,820	−12,333	−7,366
5	−13,182	−11,070	−9,481
6	−11,091	−11,497	−10,262
7	−11,307	−12,764	−8,276
8	−10,880	−10,638	−6,907
9	−13,249	−12,942	−10,126
10	−8,028	−13,921	−11,186
11	−10,930	−10,353	−11,150
12	−13,360	−10,355	−8,318
13	−11,183	−11,016	−9,250
14	−10,421	−12,345	−10,756
15	−10,949	−11,516	−7,701
16	−11,401	−10,100	−8,286
17	−11,041	−13,272	−8,551
18	−10,196	−9,765	−8,317
19	−12,849	−10,532	−8,089
20	−11,530	−12,879	−10,769
21	−12,241	−12,352	−9,320
22	−9,591	−11,286	−8,191
23	−9,787	−11,496	−7,780
24	−11,119	−12,042	−8,094
25	−9,632	−12,134	−9,924
26	−10,527	−12,603	−9,051
27	−11,100	−13,831	−13,748
28	−12,613	−12,165	−8,767
29	12,993	−13,161	−7,822
Mean	−10,383	−11,892	−9,003

features of the external phthalimide moiety (ring C) of **15** with those of the RT residues (in particular P321 and K103), confirmed that *E*,*Z* isomer is the most probable DS isomer involved in interaction with RT.

CoMFA and CoMSIA analysis

To develop the 3D-QSAR analyses, **1–29** were manually aligned on the basis of the common disulfide moiety, as shown in Fig. 7.

CoMFA analysis was performed by dividing compounds **1–29** into a training set (**1**–**6**, **8**–**12**, **14**–**18**, **20**, **22**–**26**, **28**) for model generation and into a test set (**7**, **13**, **19**, **21**, **27**, **29**) for model validation. CoMFA and CoMSIA studies were developed using CoMFA steric and electrostatic fields and CoMSIA steric, electrostatic, hydrophobic and H-bond acceptor properties, respectively, as independent variables, and the ligand pEC_{50} as the dependent variable. For CoMSIA analysis, the H-bond donor descriptor was not taken into consideration because ligands display no H-bond donor groups.

The final CoMFA model was generated using non-cross-validated PLS analysis with the optimum number of components (ONC=4) to give a non-cross validated $r^2(r^2_{ncv}) = 0.95$, standard error of estimate (SEE) = 0.264, steric contribution = 0.464 and electrostatic contribution = 0.536. The model reliability thus generated was supported by bootstrapping results. All statistical parameters supporting the CoMFA model are listed in Table 3.

A CoMSIA model consisting of steric, electrostatic, hydrophobic and H-bond acceptor fields with a $r^2_{ncv} = 0.93$, SEE=0.301, steric contribution = 0.072, electrostatic contribution = 0.462, hydrophobic contribution = 0.210, and H-bond acceptor contribution = 0.256 was derived. All statistical parameters supporting CoMSIA model are listed in Table 4.

Experimental and predicted binding affinities values for the training set and test set are reported in Table 5, while distribution of experimental and predicted pEC_{50} values for the training set and test set according to the CoMFA and CoMSIA models are shown in Fig. 8.

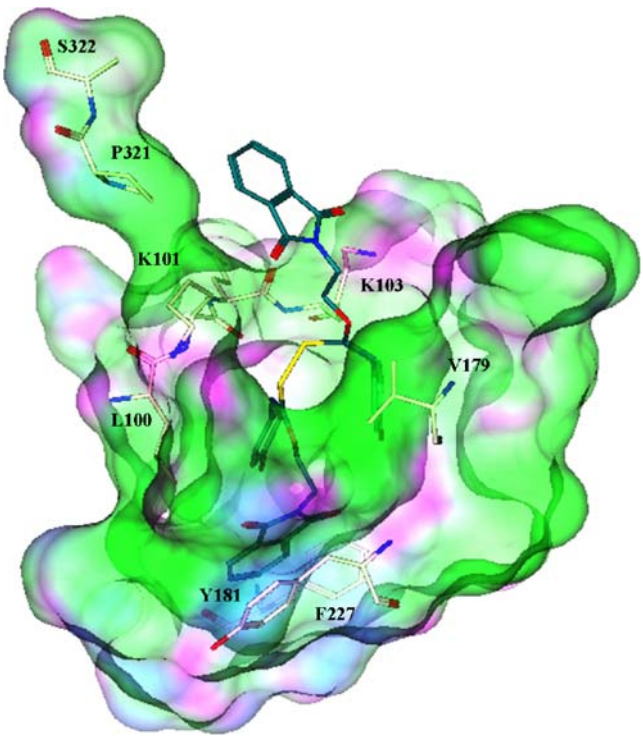


Fig. 6 Selected docking pose of **15** (*E*,*Z* isomer) into the NNIBS. The RT surface electrostatic distribution (Connolly surface) is shown. *Green areas* are related to hydrophobic regions while *magenta areas* indicate H-bond regions

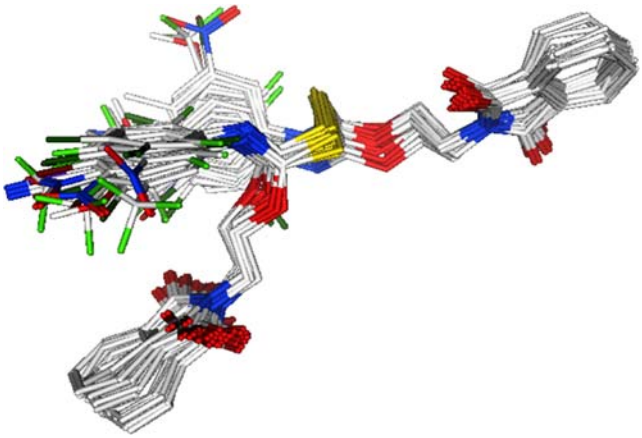


Fig. 7 Alignment of **1**–**29** used in comparative molecular fields analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA)

On the basis that CoMFA and CoMSIA field effects on the target properties can be viewed as 3D coefficient contour plots, identifying important regions where any change in these fields may affect biological activity, they could be helpful in optimising disulfides as NNRTIs. The 3D-QSAR analysis maps are described and discussed in the following sections.

CoMFA steric and electrostatic regions

As shown in Fig. 9, the steric contour map predicts favourable interaction polyhedra (green) for the 3 and 4 positions of both phthalimide moieties and the *para* position of the two *N*-phenyl rings. For all the compounds, one oxygen of A and C, and the *ortho* position of the B and D rings, are surrounded by yellow polyhedra (disfavoured). The reliability of the steric map calculations

Table 3 Summary of comparative molecular fields analysis (CoMFA) results

Number of compounds	23
Optimal number of components (ONC)	4
Leave one out $r^2\left(r^2_{loo}\right)$	0.337
Cross-validated $r_2\left(r^2_{cv}\right)$	0.482
Standard error of estimate (SEE)	0.264
Non cross-validated $r^2\left(r^2_{ncv}\right)$	0.95
F value	80.971
Steric contribution	0.464
Electrostatic contribution	0.536
Bootstrap $r_2\left(r^2_{boot}\right)$	0.97
Standard Error of Estimate r^2_{boot} (SEE r^2_{boot})	0.210
Test set $r^2\left(r^2_{pred}\right)$	0.73

Table 4 Summary of comparative molecular similarity indices analysis (CoMSIA) results

Number of compounds	23
Optimal number of components (ONC)	4
Leave one out $r^2\left(r^2_{loo}\right)$	0.391
Cross-validated $r^2\left(r^2_{cv}\right)$	0.492
Standard error of estimate (SEE)	0.301
Non cross-validated $r^2\left(r^2_{ncv}\right)$	0.93
F value	61.170
Steric contribution	0.072
Electrostatic contribution	0.462
Hydrophobic contribution	0.210
H-bond acceptor contribution	0.256
Bootstrap $r^2\left(r^2_{boot}\right)$	0.94
Standard Error of Estimate r^2_{boot} (SEE r^2_{boot})	0.262
Test set $r^2\left(r^2_{pred}\right)$	0.71

Table 5 Experimental and predicted pEC_{50} values of dataset compounds

Compound	Expected pEC_{50}	CoMFA model		CoMSIA model	
		Predicted pEC_{50}	Residual	Predicted pEC_{50}	Residual
1	6.46	6.72	−0.26	6.63	−0.17
2	5.92	5.90	0.02	5.77	0.15
3	5.64	5.60	0.04	5.42	0.22
4	5.30	5.63	−0.33	5.72	−0.42
5	6.40	5.72	0.68	5.81	0.59
6	5.22	5.74	−0.52	5.72	−0.50
7[a]	6.15	5.43	0.72	5.70	0.45
8	5.89	5.88	0.01	5.98	−0.09
9	7.40	7.35	0.05	7.32	0.08
10	7.00	6.96	0.04	7.16	−0.16
11	7.00	7.00	0.00	7.17	−0.17
12	4.62	4.85	−0.23	4.74	−0.12
13[a]	7.70	7.59	0.11	7.21	0.49
14	7.70	7.83	−0.13	7.52	0.18
15	8.00	7.72	0.28	7.37	0.63
16	7.52	7.71	−0.19	7.55	−0.03
17	7.15	7.06	0.09	7.46	−0.31
18	7.22	7.11	0.11	7.41	−0.19
19[a]	5.52	5.72	−0.20	6.18	−0.66
20	6.30	6.50	−0.20	6.22	0.08
21[a]	5.89	6.14	−0.25	5.85	0.04
22	4.74	4.69	0.05	4.53	0.21
23	5.40	5.42	−0.01	5.49	−0.09
24	5.15	5.15	0.00	5.13	0.02
25	6.70	6.57	0.14	6.50	0.20
26	4.55	4.45	0.11	4.55	0.00
27[a]	7.15	7.34	−0.19	6.73	0.42
28	5.74	5.48	0.26	5.86	−0.12
29[a]	5.92	6.50	−0.58	5.97	−0.05

[a] Test set compounds

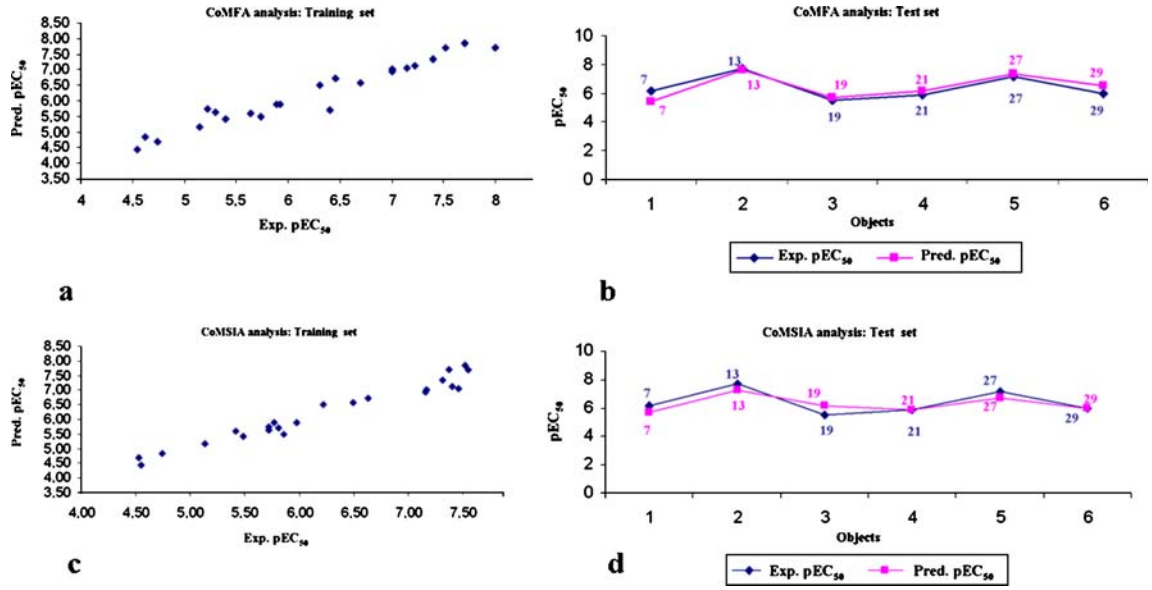


Fig. 8 Distribution of experimental and predicted pEC_{50} values for training set compounds according to CoMFA analysis (**a**), for test set compounds according to CoMFA analysis (**b**), for training set compounds according to CoMSIA analysis (**c**), and for test set compounds according to CoMSIA analysis (**d**)

is verified by the activity trend of the *N-para*-halophenyl derivatives [cf. fluoro **13** (pEC_{50}=7.70), chloro **14** (pEC_{50}=7.70), bromo **15** (pEC_{50}=8.00), iodo **16** (pEC_{50}=7.52) and unsubstituted phenyl **1** (pEC_{50}=6.46)] and by the lower pEC_{50} values of the *N-ortho*-substituted phenyl DSs [*ortho*-fluoro **2** (pEC_{50}=5.92) vs unsubstituted **1**; 2,4,6-trifluoro **29** (pEC_{50}=5.92) vs 4-fluoro **13**; 2,4-dichloro **23** (pEC_{50}=5.40) vs 4-chloro **14**; 4-bromo-2-methyl **28** (pEC_{50}=5.74) vs 4-bromo **15**]. Moreover, the presence of a small substituent is predicted to be favoured on the methylene adjacent to the phthalimide A nitrogen.

According to the electrostatic fields contour map of the CoMFA analysis plotted in Fig. 10, less positive moieties are predicted to be favoured (red areas) in the proximity of one oxygen of C, in the *meta* and *para* positions of ring D and around the *para* position of ring B. These results are in agreement with the high pEC_{50} values of the *N-para*-halo- (**13-16;** pEC_{50}=7.52-8.00), *-para*-nitro- (**17**, pEC_{50}=7.15) and -4-chloro-3-nitro-phenyl (**27**, pEC_{50}=7.15) derivatives. On the other hand, more electropositive substituents are predicted to be beneficial (blue polyhedra) around positions 3 and 4 of phthalimide A, the *ortho* and *meta* positions of phenyl D and around the *meta* position of phenyl B.

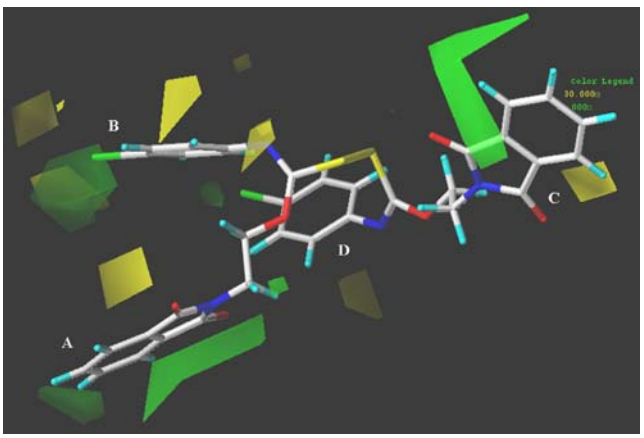


Fig. 9 Contour maps of CoMFA steric regions (*green* favoured, *yellow* disfavoured) are shown around compound **13**, reported in *stick mode* and coloured by atom type

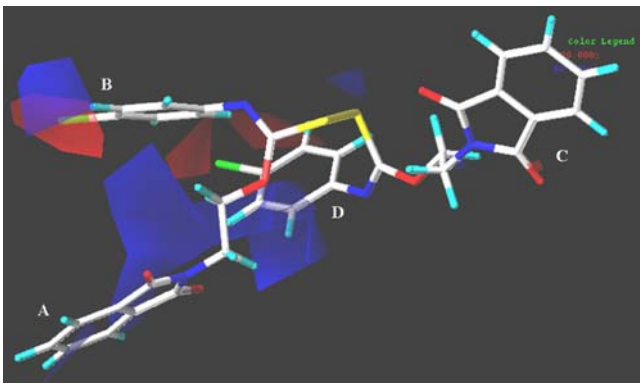


Fig. 10 Contour maps of CoMFA electrostatic regions are shown around compounds **13**. *Blue regions* are favourable for more positively charged groups; *red regions* are favourable for less positively charged groups

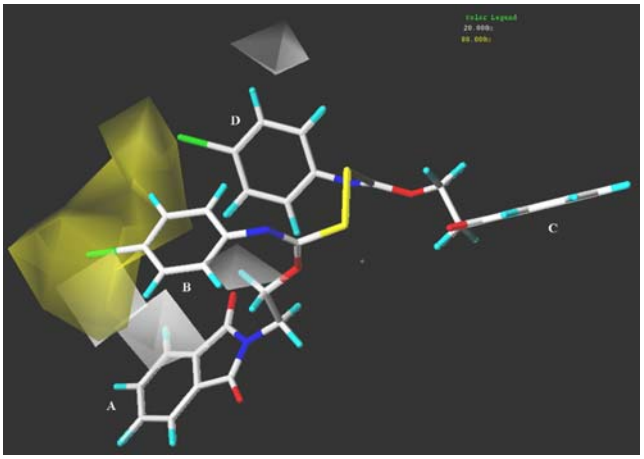


Fig. 11 Contour maps of the CoMSIA hydrophobic regions (*yellow* favoured, *white* disfavoured) are shown around compounds **13**, reported in *stick mode* and coloured by atom type

The CoMSIA steric and electrostatic regions are in agreement with the CoMFA steric and electrostatic areas.

CoMSIA hydrophobic and H-bond acceptor regions

The calculated CoMSIA hydrophobic contours (Fig. 11) predict favourable hydrophobic substituents (yellow areas) around the *para* position of rings B and D, whereas, in proximity to their *ortho* portions, lipophilic groups seem to be detrimental for activity. A *meta* lipophilic substituent is beneficial on ring B but unfavourable on ring D. These results are in agreement with the high pEC_{50} values of *N*-*para*-halo- (**13–16**) and -*para*-methylphenyl (**9**) DSs (pEC_{50}=7.40-8.00).

To take into account the role of H-bond acceptor groups for antiretroviral activity, the corresponding CoMSIA contours were calculated (Fig. 12) (CoMSIA H-bond acceptor map corresponds to the H-bond donating groups of the receptor).

As shown in Fig. 12, H-bond acceptor groups are predicted to be favoured (magenta regions) around the oxygen atoms of phatlimide A, the nitrogen atom of *N*-phenyl ring B, and in the proximity of the *para* position of ring D. In addition, H-bond acceptor functions would be unfavourable (green polyhedra) around ring B and at position 3 of phthalimide A.

The information obtained by the modelling and 3D-QSAR studies provide useful suggestions in the synthesis of DSs endowed with higher potency and an improved resistance profile. The *para* position of ring B could be exploited to establish either hydrophobic contacts with L234 and P236 (by introducing a lipophilic substituent) or an H-bond with Y318 (by introducing an H-bond acceptor function). Notably, amino acids W229, L234 and P236 are highly conserved in the NNIBS and therefore are recognised to be of strategic relevance for the design of new NNRTIs more resilient to the effects of RT mutations in this site [12]. Also, electron-positive substituents at positions 3 and 4 of phthalimide A might allow a cation–π interaction with the aromatic rings of Y181, Y188 and W229. The H-bond interaction between one of the carbonyl groups of phathalimide C (positioned outside the NNIBS pocket) and the K103 ε-amino group might allow DSs to maintain a high level of activity against the clinically relevant resistant mutant K103N (the terminal amidic group of asparagine is a good H-bond donor). Finally, the insertion of a methyl group onto the methylene adjacent to the nitrogen of phthalimide A might increase DS antiretroviral activity, as already observed in the case of C-TCs.

Conclusions

The computational studies presented here indicate the *E*,*Z* isomer as the most probable geometric isomer interacting with RT, and highlight the main interactions responsible for DS antiretroviral activity. Moreover, the present findings provide useful suggestions for the synthesis of new analogues with improved potency, also against clinically relevant resistant mutants. In the future, the models elaborated will be exploited to design new DSs and predict their activity prior to synthesis.

Acknowledgements This work was supported by University of Genoa. Progetto Ateneo 2007. Fondazione Carige is gratefully acknowledged for financially supporting E.C. and S.C.

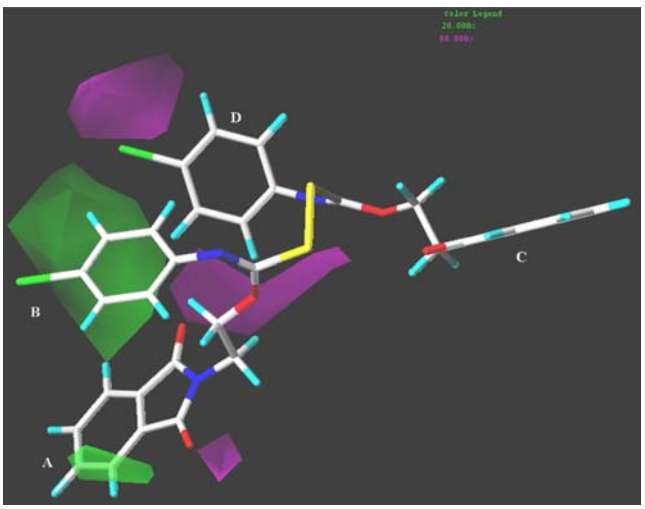


Fig. 12 CoMSIA hydrogen bond acceptor polyhedra are reported around compounds **13** depicted in *stick mode* and coloured by atom type. H-bond acceptor groups: *magenta* favoured, *green* disfavoured

References

1. Jonckheere H, Anne J, De Clercq E (2000) The HIV-1 reverse transcription (RT) process as target for RT inhibitors. Med Res Rev 20:129–154. doi:10.1002/(SICI)1098-1128(200003)20:2<129::AID-MED2>3.0.CO;2-A
2. De Clercq E (2001) New developments in anti-HIV chemotherapy. Farmaco 56:3–12. doi:10.1016/S0014-827X(01)01007-2
3. De Clercq E (2005) Emerging anti-HIV drugs. Expert Opin Emerg Drugs 10:241–273. doi:10.1517/14728214.10.2.241
4. De Clercq E (2005) New approaches toward anti-HIV chemotherapy. J Med Chem 48:1297–1313. doi:10.1021/jm040158k
5. Barbaro G, Scozzafava A, Mastrolorenzo A, Supuran CT (2005) Highly active antiretroviral therapy: current state of the art, new agents and their pharmacological interactions useful for improving therapeutic outcome. Curr Pharm Des 11:1805–1843. doi:10.2174/1381612053764869
6. Balzarini J (2004) Current status of the non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1. Curr Top Med Chem 4:921–944. doi:10.2174/1568026043388420
7. De Clercq E (1998) The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. Antivir Res 38:153–179. doi:10.1016/S0166-3542(98)00025-4
8. Pedersen OS, Pedersen EB (1999) Non-nucleoside reverse transcriptase inhibitors, the NNRTI boom. Antivir Chem Chemother 10:285–314
9. Pedersen OS, Pedersen EB (2000) The flourishing syntheses of non-nucleoside reverse transcriptase inhibitors. Synthesis 4:479–495. doi:10.1055/s-2000-6357
10. Campiani G, Ramunno A, Maga G, Nacci V, Fattorusso C, Catalanotti B et al (2002) Non-nucleoside HIV-1 reverse transcriptase (RT) inhibitors: past, present, and future perspectives. Curr Pharm Des 8:615–657. doi:10.2174/1381612024607207
11. De Clercq E (2004) Non-nucleoside reverse transcriptase inhibitors (NNRTIs): past, present, and future. Chem Biodivers 1:44–64. doi:10.1002/cbdv.200490012
12. Pauwels R (2004) New non-nucleoside reverse transcriptase inhibitors (NNRTIs) in development for the treatment of HIV infections. Curr Opin Pharmacol 4:437–446. doi:10.1016/j.coph.2004.07.005
13. Leigh Brown AJ, Frost SD, Mathews WC, Dawson K, Hellmann NS, Daar ES et al (2003) Transmission fitness of drug-resistant human immunodeficiency virus and the prevalence of resistance in the antiretroviral-treated population. J Infect Dis 187:683–686. doi:10.1086/367989
14. Richman DD, Morton SC, Wrin T, Hellmann N, Berry S, Shapiro MF et al (2004) The prevalence of antiretroviral drug resistance in the United States. AIDS 18:1393–1401. doi:10.1097/01.aids.0000131310.52526.c7
15. Spallarossa A; Cesarini S; Ranise A; Bruno O; Schenone S; La Colla P; Collu G; Sanna G; Secci B; Loddo R (2008)Novel modifications in the series of O-(2-phenylethyl)-N-substituted thiocarbamates and their ring-opened congeners as non-nucleoside HIV-1 reverse transcriptase inhibitors. Eur J Med Chem (in press). doi:10.1016/j.ejmech.2008.09.024
16. Ranise A, Spallarossa A, Cesarini S, Bondavalli F, Schenone S, Bruno O et al (2005) Structure-based design, parallel synthesis, structure-activity relationship, and molecular modeling studies of thiocarbamates, new potent non-nucleoside HIV-1 reverse transcriptase inhibitor isosteres of phenethylthiazolylthiourea derivatives. J Med Chem 48:3858–3873. doi:10.1021/jm049252r
17. Spallarossa A, Cesarini S, Ranise A, Schenone S, Bruno O, Borassi A, La Colla P, Pezzullo M, Sanna G, Collu G, Secci B, Loddo R (2008) Parallel synthesis, molecular modelling and further structure-activity relationship studies of new acylthiocarbamates as potent non-nucleoside HIV-1 reverse transcriptase inhibitors. Eur J Med Chem (in press). doi:10.1016/j.ejmech.2008.10.032
18. Ranise A, Spallarossa A, Schenone S, Bruno O, Bondavalli F, Vargiu L, Marceddu T, Mura M, La Colla P, Pani A (2003) Design, synthesis, SAR, and molecular modeling studies of acylthiocarbamates: a novel series of potent non-nucleoside HIV-1 reverse transcriptase inhibitors structurally related to phenethylthiazolylthiourea derivatives. J Med Chem 46:768–781. doi:10.1021/jm0209984
19. Cesarini S, Spallarossa A, Ranise A, Bruno O, La Colla P, Secci B et al (2008) Thiocarbamates as non-nucleoside HIV-1 reverse transcriptase inhibitors. Part 2: Parallel synthesis, molecular modelling and structure-activity relationship studies on analogues of O-(2-phenylethyl)-N-phenylthiocarbamate. Bioorg Med Chem 16:4173–4185. doi:10.1016/j.bmc.2007.12.046
20. Cesarini S, Spallarossa A, Ranise A, Fossa P, La Colla P, Sanna G et al (2008) Thiocarbamates as non-nucleoside HIV-1 reverse transcriptase inhibitors. Part 1: Parallel synthesis, molecular modelling and structure-activity relationship studies on O-[2-(hetero)arylethyl]-N-phenylthiocarbamates. Bioorg Med Chem 16:4160–4172. doi:10.1016/j.bmc.2007.12.050
21. Cesarini S, Spallarossa A, Ranise A, Schenone S, Bruno O, La Colla P et al (2008) Parallel one-pot synthesis and structure-activity relationship study of symmetric formimidoester disulfides as a novel class of potent non-nucleoside HIV-1 reverse transcriptase inhibitors. Bioorg Med Chem 16:6353–6363. doi:10.1016/j.bmc.2008.05.010
22. Spallarossa A, Cesarini S, Ranise A, Ponassi M, Unge T, Bolognesi M (2008) Crystal structures of HIV-1 reverse transcriptase complexes with thiocarbamate non-nucleoside inhibitors. Biochem Biophys Res Commun 365:764–770. doi:10.1016/j.bbrc.2007.11.036
23. MOE Chemical Computing Group Inc. Montreal. H3A 2R7 Canada. http://www.chemcomp.com
24. Cramer RD III, Patterson DE, Bunce JD (1989) Recent advances in comparative molecular field analysis. CoMFA. Prog Clin Biol Res 291:1–165
25. Klebe G, Abraham U, Mietzner T (1994) Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. J Med Chem 37:4130–4146. doi:10.1021/jm00050a010
26. Sybyl7.0, Tripos Inc 1699 South Hanley Road, St Louis, Mo, USA